Heuristically Optimized Tradeoffs

Power Laws in graphs

Kimmo Palin

Department Of Computer Science

University of Helsinki

Introduction

- It has been observed that the degree distribution of the Internet graph follows *power law*. i.e. For node i, $\Pr(\deg(i) > D) = cD^{-\beta}$.
- This distribution has strong tails. i.e. There are relatively few nodes for which there is very high degrees.
- Same property has been observed also for example in metabolic and gene networks, income distribution, city populations, word frequencies and computer file sizes.
- Power laws are known to show up in situations that require reliability but the redundancy has costs. (Highly Optimized Tolerance)
- Next we show that power laws can also come up in locally optimized tradeoffs.

Internet model

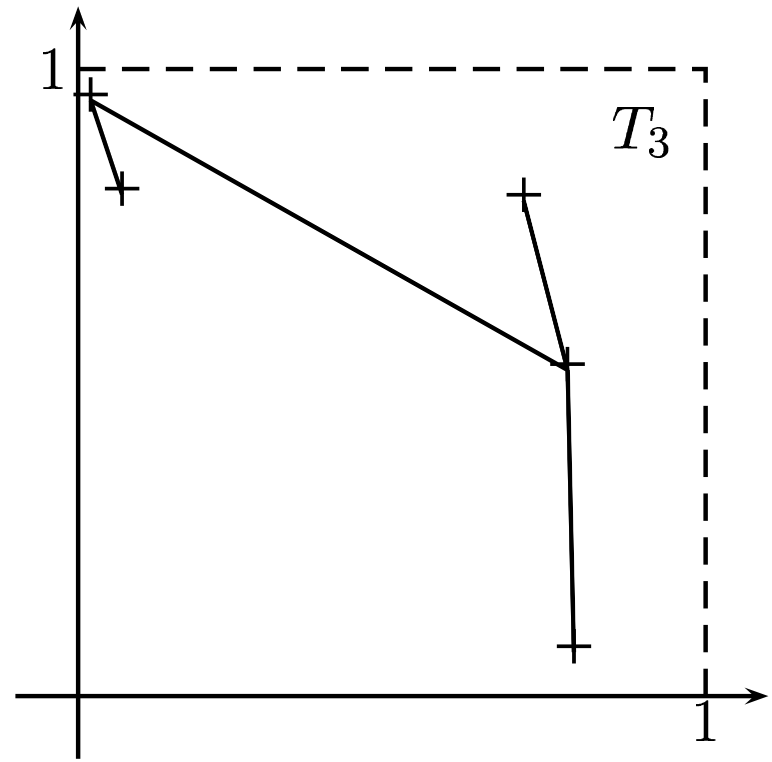
- Following model and theorem is by Alex Fabrikant, Elias Koutsoupias, Christos Papadimitriou: Heuristically Optimized Tradeoffs: A New Paradigm for Power Laws in the Internet
- After the early begining there hasn't been much need for extra robustness in the internet, still the degrees are distributed as power law.
- Consider Internet as nondirected tree where each node, first coming to the Net, wants to be (topologically) close to the "center" of the network but has to pay for the cable between itself and the center.
- In reality one might want to minimize the average or maximum hop–count instead of hops to the center.
- Nodes coming on–line define sequence of internet trees T_i.

Sequence of trees T_i

- Sequence of nodes (points) $p_0, p_1, \ldots, p_n$ arrive uniformly at random to a unit square.
- New node p_i connects to some previous node by minimizing the joint cost of the "last mile" and number of hops. That is, to p_j that minimizes

$$f_i(j) = \alpha \cdot d(i,j) + h(j)$$

- $d(i,j)$ is the euclidean distance between p_i and p_j and $h(j)$ is the graph distance from p_j to p_0.

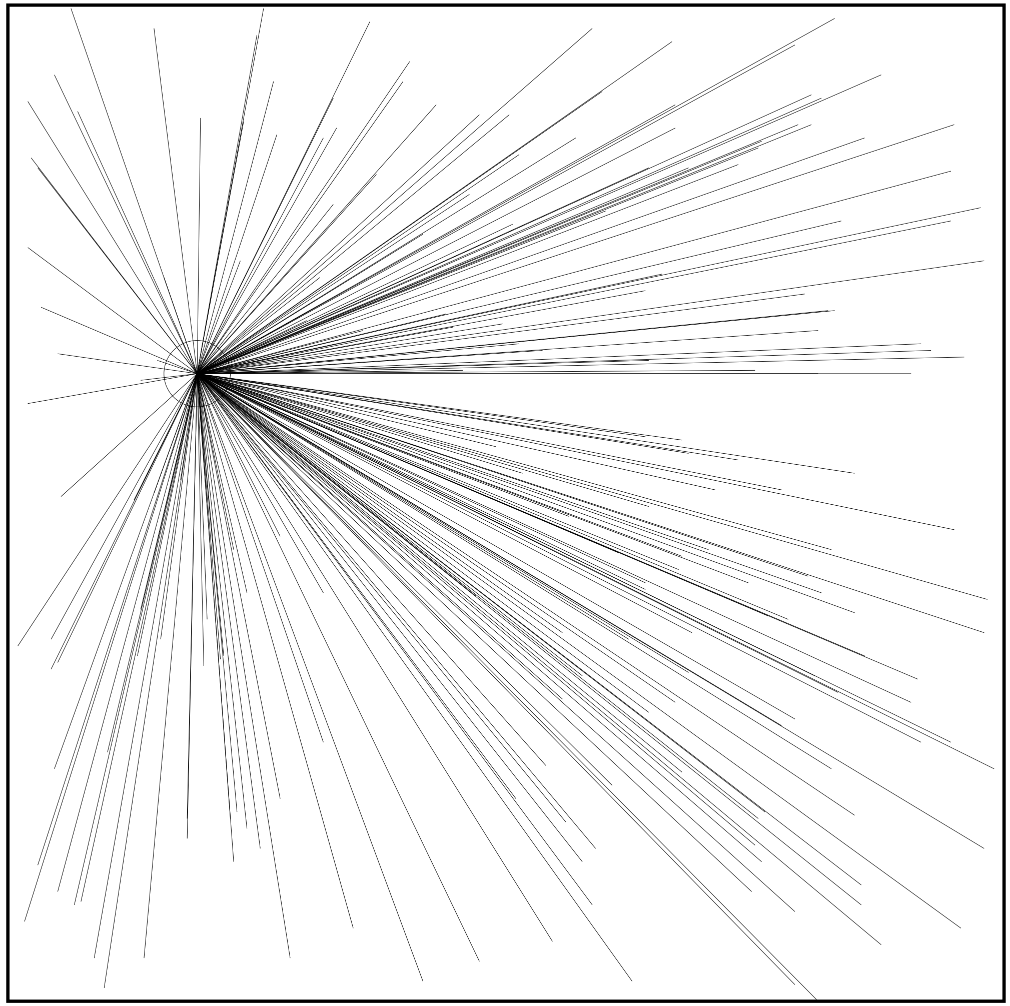


Theorem

Theorem If T_n is generated as above, then:

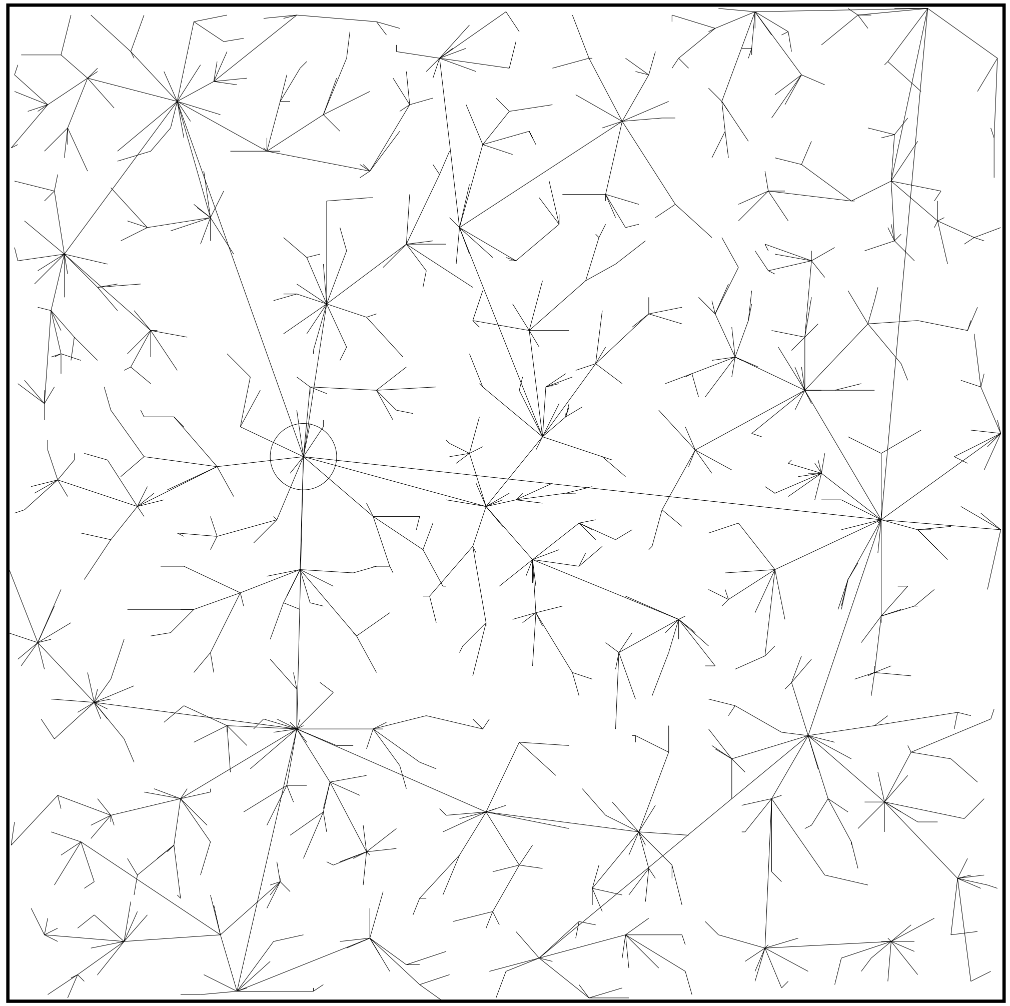
1. If $\alpha < 1/\sqrt{2}$, then T_n is a star with p_0 as its center.
2. If $\alpha = \Omega(\sqrt{n})$, then the degree distribution of T_n is exponential, that is, the expected number of nodes that have degree at least D is at most $n^2 e^{-cD}$ for some constant c.
3. If $\alpha \geq 4$ and $\alpha = o(\sqrt{n})$, then the degree distribution of T_n is a power law. i.e. the expected number of nodes with degree at least D is greater than $c \cdot (D/n)^{-\beta}$ for some constants c and β.

Case 1 $\alpha < 1/\sqrt{2}$



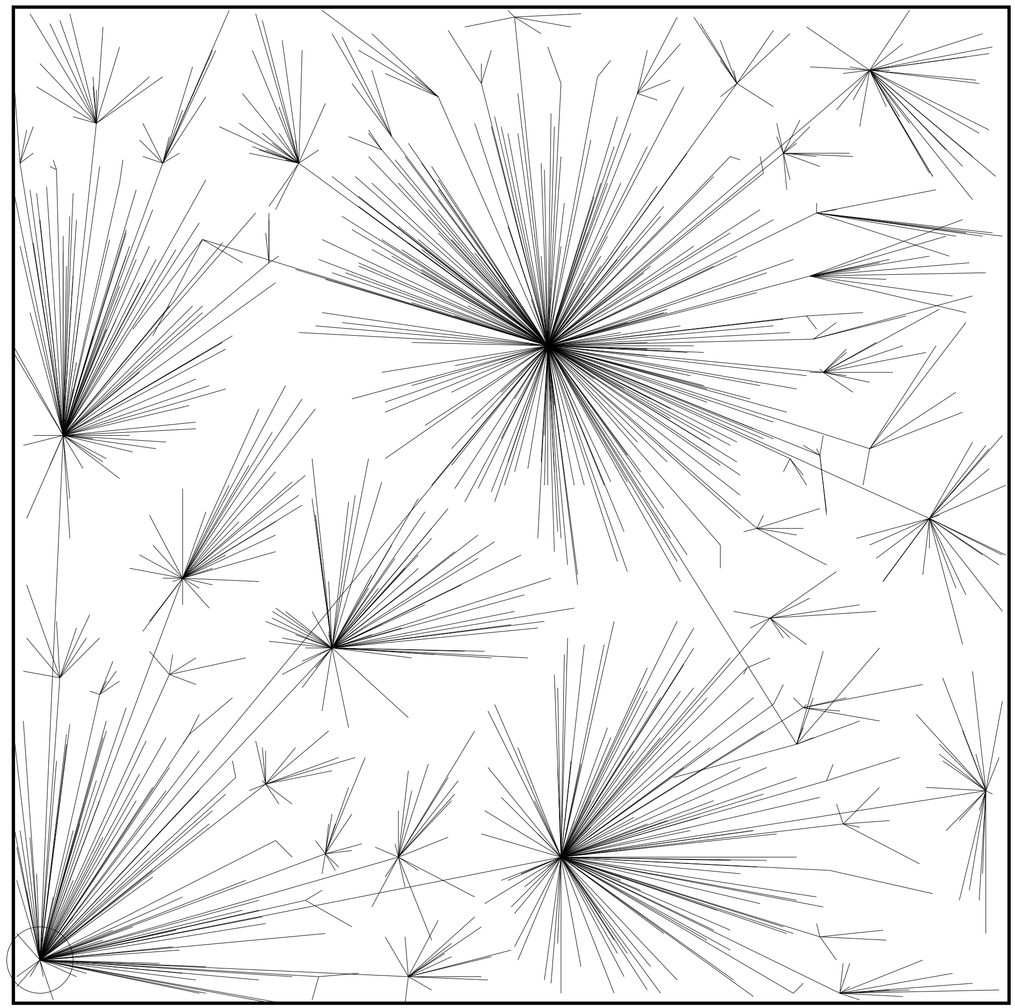
$n = 200$ ja $\alpha = 1/\sqrt{2}$

Case 2 $\alpha = \Omega(\sqrt{n})$



$n = 1000$ ja $\alpha = 40$

Case 3 $4 \leq \alpha = o(\sqrt{n})$



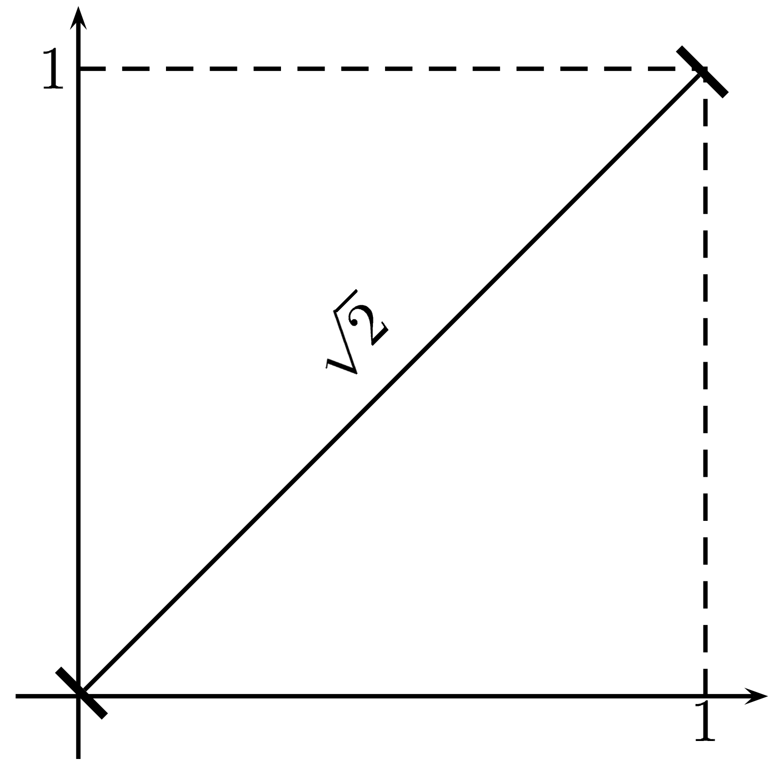
$n = 1000$ ja $\alpha = 4$

Case 1 $\alpha < 1/\sqrt{2}$

- Maximum distance between any two nodes is $\sqrt{2}$. Especially $d(i,0) < \sqrt{2}$ for all i.
- For incoming node p_i function $f_i(j)$ is

$$f_i(0) < 1/\sqrt{2} \cdot d(i,0) + h(0) \leq 1 \leq f_i(j)$$

  where $j \neq 0$.
- Thus, each incoming node p_i connects straight to p_0 and the topology of T_n is p_0 centered star. $\square$



Case 2 $\alpha = \Omega(\sqrt{n})$ I

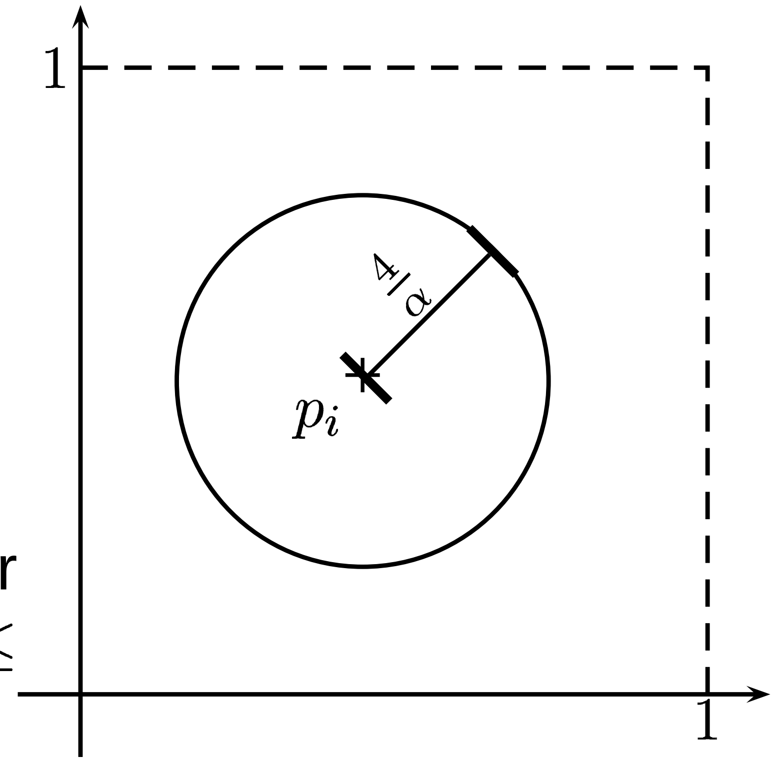
- Let neighborhood $N_k(i) = \{j | [i,j] \in T_k\}$.
- Short links $S_k(i) = |\{j \in N_k(i) | d(i,j) \leq \frac{4}{\alpha}\}|$.
- Long links $L_k(i) = |\{j \in N_k(i) | d(i,j) > \frac{4}{\alpha}\}|$.
- By the union bound

$$\Pr(\deg(i) \geq D) \leq \Pr(S_n(i) \geq D/2) + \Pr(L_n(i) \geq D/2)$$

.

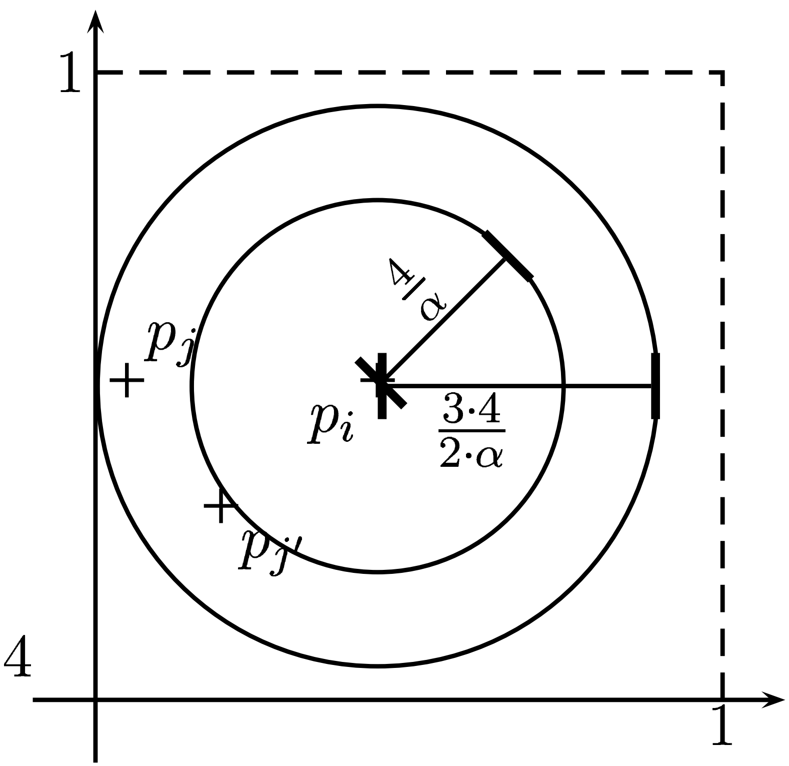
Case 2 $\alpha = \Omega(\sqrt{n})$ II

- Nodes contributing to $S_n(i)$ have to fall in to circle of area $\pi(\frac{4}{\alpha})^2 \leq 16\pi n^{-1}$
- Because of i.i.d new nodes, $S_n(i)$ is bounded by sum of Bernoulli trials with $\mathrm{E}S(i) = \Theta(n\alpha^{-2}) < c$.
- By the Chernoff–Hœffding bound, for $D > 3c$, $\Pr(S(i) > D/2) \leq \exp(-\frac{(D-2c)^2}{D+4c}) \leq e^{-D/21}$.



Case 2 $\alpha = \Omega(\sqrt{n})$ III

- Let $L_x(i) = |\{j \in N_n(i) | x \leq d(i,j) \leq \frac{3}{2}x\}|$.
- Show that $L_x(i) < 14$ for all i and $x \geq \frac{4}{\alpha}$.
- Let points p_j and $p_{j'}$, $j \leq j'$ lie between x and $\frac{3}{2}x$ away from p_i.
- If $|\angle p_j p_i p_{j'}| < c = \cos^{-1}(43/48)$, then $p_{j'}$ prefers p_j over p_i, since $h(j) - h(i) \leq 1$ and $\alpha d(i,j') > \alpha d(j,j') + 1$.
- Because $\cos^{-1}(43/48) > \frac{2\pi}{14}$, $L_x(i) < 14$ for all $x \geq \frac{4}{\alpha}$.



Case 2 $\alpha = \Omega(\sqrt{n})$ IV

- Let $\delta_i = \max\{\frac{4}{\alpha}, \min_j \ d(i,j)\}$ and bind $L(i)$ by

$$L(i) = \sum_{k=1}^{-\log_{\frac{3}{2}} \delta_i} L_{(\frac{3}{2})^{-k}}(i) \leq -14 \log_{\frac{3}{2}} \delta_i$$

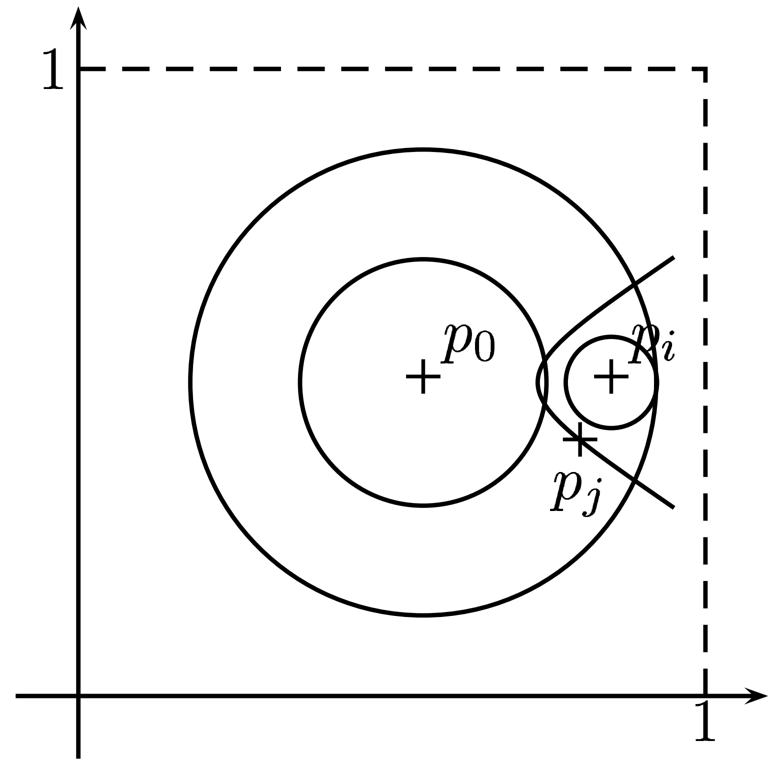
- Since points are distributed uniformly and independently, $\Pr(\delta_i \leq y) \leq 1 - (1 - \pi y^2)^{n-1} \leq \pi(n-1)y^2$.
- Therefore

$$\Pr(L(i) \geq D/2) \leq \Pr(-14 \log_{\frac{3}{2}} \delta_i \geq D/2) \leq \pi(n-1)(\frac{3}{2})^{D/14}$$

$\square$

Case 3 $4 \le \alpha = o(\sqrt{n})$

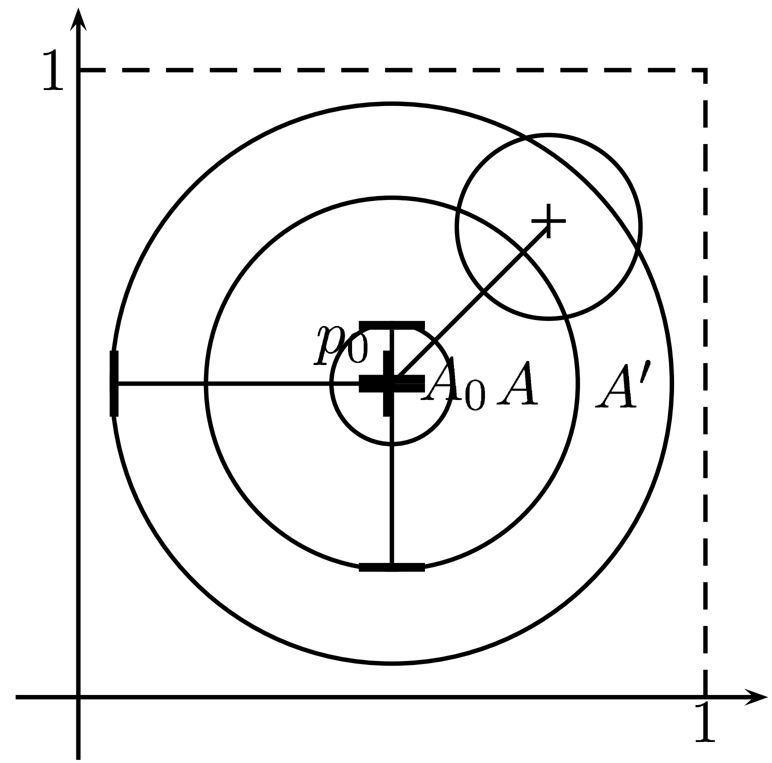
- Let $r(i) = d(i,0) - \frac{1}{\alpha}$. If $r(i) < 0$, then i will connect to 0. Next consider p_i such that $\frac{1}{\alpha} < d(i,0) < \frac{3}{2\alpha}$ and connects to p_0.



Lemma 1: Every point arriving after p_i inside the circle of radius $\frac{1}{4}r(i)$ around p_i will link to p_i.
Lemma 2: No point p_j will link to p_i unless $|\angle p_j p_0 p_i| \le \sqrt{2.5\alpha r(i)}$ and $d(j,0) \ge \frac{1}{2}r(i) + \frac{1}{\alpha}$.

Case 3 $4 \le \alpha = o(\sqrt{n})$

- For simplicity, we consider only the case $\alpha = o(\sqrt[3]{n})$ which leads to $\Pr(\deg(i) > D) = cD^{-\frac{1}{6}}$
- Let $D \le \frac{n^{1-\epsilon}}{256\alpha^3}$ and $\rho = 4\sqrt{D/n}$ and $m = \lceil \frac{1}{2\rho} \rceil$. Now consider T_m.
- Observe p_0 centered circles A_0, A and A' of with radii in range $[\frac{1}{\alpha}, \frac{1}{\alpha} + \rho]$, $(\frac{1}{\alpha} + \rho, \frac{1}{\alpha} + \rho^{2/3}]$, $(\frac{1}{\alpha} + \rho, \frac{1}{\alpha} + 0.5\rho^{2/3}]$
- Each point in T_m that is connected to p_0 and lies in A' has region of influence of size $\pi D/n$ (lemma 1) and is expected to be linked to $\pi D/2$ nodes in T_n.



Case 3 $4 \le \alpha = o(\sqrt{n})$

- By lemma 2 and triangle inequality, all points arriving to A' connects to some point in A
- Each point that arrives to A and connects to neighbour of p_0 claims a sector no larger than $\sqrt{10\alpha}\rho^{1/3}$ of A' for itself (lemma 2).
- There are at most $\frac{1}{8}\sqrt{\alpha}\rho^{1/3}$ such sectors.
- By Chernoff bounds it is possible to show that in these sectors there are enough points with expected degree of $\pi D/2$ in T_n. (With exponentially high probability)

$\ldots$ for $C = \frac{1}{2^{16/33} n^{5/6} \sqrt{\alpha}}$, and any D in the above specified range, the probability that a randomly chosen point in T_n has degree at least D is at least $\ldots\ CD^{-1/6}$.